農研機構
NARO 国立研究開発法人 農業・食品産業技術総合研究機構

Crystal structure and substrate recognition mechanism of Aspergillus oryzae isoprimeverose-producing enzyme

メタデータ	言語: eng 出版者: 公開日: 2019-10-04 キーワード (Ja): キーワード (En): 作成者: 藤本, 瑞 メールアドレス: 所属:
URL	https://repository.naro.go.jp/records/2893

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 International License.

Crystal structure and substrate recognition mechanism of *Aspergillus oryzae* isoprimeverose-producing enzyme

Tomohiko Matsuzawa[1#], Masahiro Watanabe[2#], Yusuke Nakamichi[2], Zui Fujimoto[3], Katsuro Yaoi[1*]

[1]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

[2]Research Institute for Sustainable Chemistry, National Institute of Advanced Industrial Science and Technology (AIST), 3-11-32, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-0046, Japan

[3]Advanced Analysis Center, National Agriculture and Food Research Organization (NARO), 2-1-2 Kannondai, Tsukuba, Ibaraki, 305-8602, Japan

*Corresponding author:

K. Yaoi

Phone: +81 29 861 7867

Fax: +81 29 861 6226

E-mail address: k-yaoi@aist.go.jp

#Contributed equally

Summary

Aspergillus oryzae isoprimeverose-producing enzyme (IpeA) releases isoprimeverose (α-D-xylopyranosyl-(1→6)-D-glucopyranose) from the non-reducing end of xyloglucan oligosaccharides. IpeA is classified as a member of the glycoside hydrolase family 3 (GH3). We solved the crystal structure of IpeA with isoprimeverose at 2.4 Å resolution, showing that the structure of IpeA formed a dimer and was composed of three domains: an N-terminal $(\beta/\alpha)_8$ TIM-barrel domain, α/β/α sandwich fold domain, and a C-terminal fibronectin-like domain. The catalytic TIM-barrel domain possessed a catalytic nucleophile (Asp300) and acid/base (Glu524) residues. Interestingly, we found that the cavity of the active site of IpeA was larger than that of other GH3 enzymes, and subsite −1' played an important role in its activity. The glucopyranosyl and xylopyranosyl residues of isoprimeverose were located at subsites −1 and −1', respectively. Our findings provide new insights into the substrate recognition of GH3 enzymes.

Keywords

Xyloglucan, glycoside hydrolase, isoprimeverose, *Aspergillus oryzae*, glycoside hydrolase family 3, hemicellulose, oligosaccharide

Introduction

Xyloglucan, one of the most widely distributed polysaccharides in the primary cell wall and seeds of plants, plays an important role in plant growth and development (Carpita et al., 1993). Xyloglucan consists of a β-1,4-glucan backbone and α-linked xylopyranosyl residues attached to the C6 position of the glucopyranosyl residues of the β-1,4-glucan backbone. Some branched xylopyranosyl residues are modified with β-1,2-linked galactopyranosyl residues. Additionally, xyloglucan is modified with other saccharides, including arabinose and fucose (Zabotina, 2012). The structures of xyloglucan and xyloglucan oligosaccharides are indicated with abbreviations, such as G, unbranched glucopyranosyl residue; X, α-D-xylopyranosyl-(1→6)-β-D-glucopyranosyl segment; and L, β-D-galactopyranosyl-(1→2)-α-D-xylopyranosyl-(1→6)-β-D-glucopyranosyl segment (Fry et al., 1993; Tuomivaara et al., 2015). Xyloglucan is degraded by many combinations of glycoside hydrolases, such as xyloglucan-specific endo-glucanases (xyloglucanase, EC 3.2.1.151) (Edwards et al., 1986; Grishutin et al., 2004; Matsuzawa et al., 2014; Yaoi and Mitsuishi, 2004; Yaoi et al., 2005), oligoxyloglucan-reducing end-specific cellobiohydrolases (EC 3.2.1.150) (Bauer et al., 2005; Yaoi et al., 2002), α-xylosidases (EC 3.2.1.177) (Larsbrink et al., 2014; Matsuzawa et al., 2016b; Moracci et al., 2000; Okuyama et al., 2004), β-galactosidases (EC 3.2.1.23) (Edward et al., 1988; York et al., 1993), and isoprimeverose-producing enzymes (EC 3.2.1.120) (Kato et al., 1985; Matsuzawa et al., 2016c; Yaoi et al., 2007; Yaoi and Miyazaki, 2012). To date, only two isoprimeverose-producing enzymes have been isolated and characterized. Bacterial and fungal isoprimeverose-producing enzymes were isolated from *Oerskovia* sp. (Yaoi et al., 2007; Yaoi and Miyazaki, 2012) and *Aspergillus oryzae* (Kato et al., 1985; Matsuzawa et al., 2016c). *A. oryzae* isoprimeverose-producing enzyme (IpeA) strictly recognizes and releases isoprimeverose units from the non-reducing end of xyloglucan oligosaccharides. For example, IpeA hydrolyzes XG

61 (trisaccharide, Xyl_1Glc_2) to X (isoprimeverose, Xyl_1Glc_1) and D-glucose, and XXXG (heptasaccharide,
62 Xyl_3Glc_4) to X and XXG (Xyl_2Glc_3) (Matsuzawa et al., 2016c). IpeA prefers XXXG to XG and XX
63 (Xyl_2Glc_2), suggesting that IpeA has more than three subsites (−1, +1, and +2). Galactosylation at the
64 α-1,6-linked xylopyranosyl side chains of the non-reducing ends of xyloglucan oligosaccharides
65 completely abolishes IpeA activity. IpeA not only exhibits hydrolytic activity toward xyloglucan
66 oligosaccharides but also transglycosylation activity, and can produce many kinds of oligosaccharides
67 (Matsuzawa et al., 2016c).

68 Glycoside hydrolases are divided into glycoside hydrolase families (GH) based on their amino acid
69 sequences (carbohydrate-active enzymes; CAZy database, http://www.cazy.org/) (Lombard et al., 2013).
70 Isoprimeverose-producing enzymes belong to GH3. Examples of GH3 enzymes include β-glucosidases
71 (EC 3.2.1.21), β-xylosidases (EC 3.2.1.37), β-glucosylceramidase (EC 3.2.1.45),
72 isoprimeverose-producing enzymes (EC 3.2.1.120), and so on. Previously, GH3 enzymes were thought to
73 recognize and release monosaccharides from the non-reducing end. However, isoprimeverose-producing
74 enzymes recognize and release disaccharide units, isoprimeverose
75 (α-D-xylopyranosyl-(1→6)-D-glucopyranose), from the non-reducing end of xyloglucan oligosaccharides
76 (Matsuzawa et al., 2016c; Yaoi et al., 2007), indicating that the substrate recognition machinery of
77 isoprimeverose-producing enzymes differs from that of other GH3 enzymes.

78 Although isoprimeverose-producing enzymes have unique substrate specificity and transglycosylation
79 activity, the structure of isoprimeverose-producing enzymes remains unclear. In this study, we solved the
80 crystal structure of *A. oryzae* isoprimeverose-producing enzyme (IpeA). IpeA had an isoprimeverose
81 recognition pocket (subsites −1 and −1′) and recognized β-glucan main chain and branched xylose at both

82 the negative and positive subsites. Our findings clearly demonstrate the mechanism by which
83 isoprimeverose-producing enzymes recognize xyloglucan oligosaccharides.
84
85 **Results and Discussion**
86 **Overall structure of IpeA**
87 The structure of IpeA was solved using the molecular replacement method with the structure of
88 β-glucosidase JMB19063 (PDB 3U4A) as the search model. The structures of the apo form and
89 isoprimeverose complex were determined at 1.98 and 2.4 Å resolution, respectively (Table 1). Both
90 structures could be built from Lys26 to Gln779. In the crystal structure of IpeA, the asymmetric unit
91 contained two polypeptides (chains A and B) correlating with a non-crystallographic two-fold axis (Fig.
92 1A). This dimer assembly was consistent with the data obtained from gel-filtration chromatography (data
93 not shown), suggesting that the crystal structure of IpeA corresponds to the biologically relevant form of
94 the protein. Two chains A and B in the dimer had almost the same structures with with root-mean square
95 differences (rmsd) of 1.4 Å for equivalent Cα atoms. The monomeric structure of IpeA was composed of
96 three distinct domains (Fig. 1B). The N-terminal domain was a $(\beta/\alpha)_8$ barrel, the TIM-barrel domain
97 (residues 48–371), which contains the catalytic machinery and is a common structure among glycoside
98 hydrolases. The second domain was an α/β/α sandwich fold domain with central β-sheets and α-helices
99 (residues 411–660). The C-terminal domain was a fibronectin-like domain with antiparallel β-sheets
100 (residues 698–767). A domain search for the enzyme was performed using Pfam, the protein families
101 database (https://pfam.xfam.org/) (Eberhardt et al., 2016; Finn et al., 2014). These three domains of
102 chains A and B in IpeA formed extensive interactions with all of the domains of the other chain. Because
103 the fraction of buried atoms and their interactions (salt bridges, ion networks, and hydrogen bonds)

104 represent features that may be important for thermostability, we explored the dimer interface of IpeA
105 using PISA software (http://www.ebi.ac.uk/pdbe/pisa) (Krissinel and Henrick, 2007). Our analysis
106 showed that a large surface area of 2948.8 Å^2 was buried, and 42 hydrogen bonds and two salt bridges
107 were formed upon dimerization of IpeA. Both the apo and isoprimeverose complex structures of IpeA
108 were glycosylated by two species of saccharide chains, *N*-acetyl-β-D-glucosamine and chitin disaccharide
109 at asparagine residues (data not shown). Additionally, significant electron density corresponding to a
110 metal was observed at the C-terminal domain of IpeA. The metal was coordinated by the Asp706 side
111 chains, the main chain carbonyl of Val708, and four water molecules in IpeA. These distances were 2.2 to
112 2.5 Å from the metal (data not shown). Because the crystallization condition of IpeA involved the
113 reservoir solution, which includes calcium chloride (50 mM), we modeled a calcium ion for the metal in
114 the corresponding site. Previously, Kato et al. (1985) reported that the addition of calcium chloride did
115 not enhance the activity of IpeA. Therefore, the calcium ion does not appear to play an important role in
116 the enzymatic activity of IpeA.
117
118 **Structural comparison with other GH3 enzymes**
119 A structural similarity search using the Dali server (http://ekhidna2.biocenter.helsinki.fi/dali/) (Liisa and
120 Laura, 2016) toward whole protein IpeA revealed that the dimer structure of IpeA resembled other GH3
121 β-glucosidases: β-1,2-glucooligosaccharide-specific β-glucosidase Lin1840r (Z score 45.1; PDB 4zoe)
122 (Nakajima et al., 2016), metagenomic β-glucosidase JMB19063 (Z score 44.0; PDB 3u48), and barley
123 β-D-glucan exohydrolase isoenzyme ExoI (Z score 43.3; PDB 1ex1) (Varghese et al., 1999).
124 Despite low sequence similarities (22–28%), the three domains of monomer IpeA were readily
125 superimposed on GH3 β-glucosidases with Z scores 43–45 and rmsd of 1.8 to 2.0 Å for equivalent Cα

126 atoms, except for 80 residues of the C-terminal region of IpeA (Fig. 2). These results indicate that IpeA
127 belongs to a GH3 family at the level of individual domain folds and domain orientation.
128
129 **Machinery of isoprimeverose recognition**
130 The 2.4-Å resolution structure of the product (isoprimeverose) complex revealed unambiguous electron
131 density, as a disaccharide of isoprimeverose was observed in the deep pocket within the TIM-barrel
132 domain of each monomer (Fig. 3A). To provide insight into the substrate recognition of IpeA, we
133 compared the monomer structure of IpeA with the metagenomic GH3 β-glucosidase JMB19063
134 complexed with D-glucose (PDB 3U48). Interestingly, comparison of the overall structures of IpeA and
135 JMB19063 revealed differences in the cleft size of their active sites (Fig. 3A, B), although they had a high
136 degree of structural similarities. The active site of JMB19063 could only occupy a D-glucose molecule. In
137 contrast, the active site of IpeA exhibited a large enough cleft size to adopt a long-chain oligosaccharide
138 such as XXXG. Additionally, our complex structure of IpeA-isoprimeverose exhibited not only general
139 subsite −1 but also unique branching subsite −1′. In subsite −1, both the glucopyranose moiety of
140 isoprimeverose in IpeA and the D-glucose in JMB19063 sat in almost the same position that adopted a
141 chair conformation with its C1 hydroxyl group (O1) in the β-anomer configuration. These D-glucose
142 molecules formed direct hydrogen bonds with six amino acid residues of IpeA and JMB19063. These six
143 residues of IpeA, Lys220 (interacting with O3 and O4), His221 (O3), Glu125 (O4), Arg187 (O2), Asp300
144 (C1, O2), and Glu524 (O1), corresponded to Lys181 (O3), His182 (O3), Asp84 (O4, O6), Arg142 (O2),
145 Asp261 (C1, O2), and Glu488 (O1) of JMB19063, respectively, and made almost identical
146 protein-carbohydrate interactions with no apparent large differences in the inter-atomic distances (Fig. 3C,
147 D). Lys220 and His221 of IpeA formed hydrogen bonds with O3 of the glucosyl residue at 2.8 and 2.9 Å,

148 respectively, via its side-chain amide. Accordingly, the structurally equivalent residues of JMB19063
149 were Lys181 and His182, which fulfilled the same function of hydrogen bonds to O3 at 2.8 and 2.9 Å,
150 respectively (Fig. 3E, F). Glu125 of IpeA also formed a hydrogen bond with O4 (2.5 Å), but at the
151 position replaced by Asp84 in JMB19063, which formed two hydrogen bonds with O4 (2.8 Å) and O6
152 (2.8 Å). Additionally, Arg187 of IpeA formed a substitution hydrogen bond with O2 (2.9 Å), which
153 overlapped well with Arg142 (2.7 Å) of JMB19063 (Fig. 3E, F). The above interactions, which were
154 important for subsite −1 recognition of the substrate of IpeA, are conserved across GH3 β-glucosidases
155 (data not shown).
156 Glu524 of IpeA formed a hydrogen bond with O1 (2.5 Å) and was oriented toward the glycosidic
157 oxygen, whereas Asp300 formed a hydrogen bond with O2 (2.5 Å) and was suitably positioned to serve
158 as a catalytic acid/base (Fig. 3E). Consistent with the predicted roles of these residues, mutations in
159 Glu524 or Asp300 of IpeA (Glu524Ala or Asp300Ala, respectively) virtually inactivated the enzyme
160 (data not shown). In addition, barley ExoI, a GH3 member, implicated an aspartate residue (Asp285)
161 within the sequence Gly-Phe-Val-Ile-Ser-Asp-Trp in a $(\beta/\alpha)_8$ TIM-barrel domain as being appropriately
162 positioned to act as a nucleophile during catalysis (Varghese et al., 1999). Additionally, the C-terminal
163 domain contained glutamate residues (Glu491) projecting into the active site of the $(\beta/\alpha)_8$ TIM-barrel
164 domain, which was proposed to act as the catalytic acid/base. Therefore, the steric positions of Asp33 and
165 Glu524 in IpeA were consistent with the common structural features of known retaining enzymes. These
166 results strongly suggest that IpeA hydrolyzes its substrate in a double displacement retaining mechanism
167 using the acid/base Glu524 and the nucleophile Asp300, in a similar fashion to the GH3 enzymes
168 characterized thus far.

In subsite −1', although IpeA possesses the site necessary for binding a branching xylose chain, JMB19063 does not, due to the presence of sterically hindering residues of Phe54 (Fig. 4). This phenomenon is well conserved in other GH3 enzymes, ExoI and Lin1840r, which contain Leu residues at the same position (Fig. 4). Subsite −1' seemed to play a key role in correctly adjusting the position of XX (or XXXG) to hydrolyze the backbone of glucose chains. The xylose moiety of isoprimeverose made two hydrogen bonds with Gln58, and a hydrophobic stacking with Tyr89 in IpeA (Fig. 3C). The three direct substrate-binding residues were Gln58 (O2, O3), and Tyr89 (the planar face of xylose). The side-chain Nδ2 and Oδ1 of Gln58 formed two hydrogen bonds to O2 (3.0 Å) and O3 (2.7 Å) of xylose via its side-chain amide and oxygen, respectively (Fig. 3E). The aromatic side-chain of Tyr89 formed hydrophobic stacking with the planar face of xylose.

Mutations in subsite −1'

From the IpeA isoprimeverose complex structure, we speculated that the Gln58 and Tyr89 residues contributed to the recognition of a xylopyranosyl residue of isoprimeverose at the non-reducing end of xyloglucan oligosaccharides. The Tyr89 residue was replaced with an Ala, Trp, or Phe residue. The Tyr89Ala mutant completely abolished the hydrolysis activity toward XG, but some activity was detected in the cases of XX and XXXG (Table 2). The Tyr89Trp and Tyr89Phe mutants showed hydrolytic activity toward the xyloglucan oligosaccharides XG, XX, and XXXG, indicating that an aromatic residue at subsite −1' was essential for the substrate recognition and hydrolytic activity of IpeA. Mutation of the Gln58 residue located at subsite −1' completely abolished the hydrolytic activity toward XG, but activity toward XX and XXXG remained (Table 2).

190 The K_m values of the Tyr89Trp and Tyr89Phe mutants were nearly equal to that of the wild-type
191 enzyme; however, the K_m values of the Gln58Ala and Tyr89Ala mutants were higher than that of the
192 wild-type enzyme (Table 3). In addition, the k_{cat} values of the Gln58Ala and Tyr89Ala mutants were
193 dramatically decreased. These results indicate that the recognition of the xylopyranosyl residue at subsite
194 −1' using Gln58 and an aromatic residue was crucial for IpeA activity.
195 As described above, both the Tyr89 and Gln58 residues were essential for hydrolytic activity toward
196 XG, but not toward XX and XXXG (please note that these residues were not essential, but vital for the
197 hydrolysis of XX and XXXG), suggesting that the recognition of other xylopyranoside side chains of
198 xylooligosaccharides in positive subsites partially complemented the mutations in the Tyr89 and Gln58
199 residues.
200
201 **Recognition of xyloglucan oligosaccharides at positive subsites**
202 Previously, we reported that IpeA preferred XX over XG, and XXXG over XX. This result indicated that
203 IpeA recognizes branched xylopyranose residues at subsite +1' of XX and the β-1,4-glucan backbone at
204 subsite(s) +2 and/or +3 (Matsuzawa et al., 2016c). As described above, we deduced the positive subsites
205 (+1, +2, and +3) from a model of the IpeA-XXXG complex based on our IpeA-isoprimeverose structure
206 (Fig. 5A). From the model, it was predicted that at least five residues of IpeA (Glu67, Tyr92, Tyr268,
207 Arg306 and Trp515) directly contact subsites +1, +1', +2, +2', and +3 of XXXG, respectively. Notably,
208 Tyr92 and Tyr268 of IpeA are more likely to form hydrophobic stacking with the planar face of xylose at
209 subsites +1' and +2', as is the case with Tyr89 of subsite −1'. Additionally, Trp515 of IpeA was also
210 closely located at the glucose backbone at subsites +1 and +2 to contact aromatic stacking (Fig. 5B). As

211 for Arg306 and Glu67 of IpeA, these residues are thought to contribute to stabilizing the binding of
212 branching xylose chains of subsites +1' and +2', respectively, via their side-chains.
213
214 **Conclusions**
215 Through crystal structure analysis and site directed mutagenesis of IpeA, we clarified the machinery of
216 substrate recognition and hydrolytic activity of isoprimeverose-producing enzyme. Although IpeA
217 belongs to the GH3 family based on its amino acid sequence, IpeA had a larger cleft size to adopt long-
218 and branched-chain oligosaccharides, and many unique amino acid residues that contributed to substrate
219 recognition in both the negative and positive subsites. Subsite −1 of IpeA resembled those of other GH3
220 β-glucosidases, but subsite −1′ was unique and essential to isoprimeverose-producing enzymes. The
221 transglycosylation reaction of retaining glycoside hydrolases requires proper binding of an acceptor
222 saccharide after the formation of a covalent enzyme-glycosyl intermediate, and positive subsites provide a
223 structural platform for an acceptor molecule in the transglycosylation reaction (Matsuzawa et al., 2016a).
224 Because IpeA had a large cleft and many amino acid residues were involved in substrate binding in its
225 positive subsites, we speculate that these structures and amino acid residues are closely related to not only
226 its hydrolytic activity toward xyloglucan oligosaccharides but also its transglycosylation activity. Our
227 findings will contribute to a better understanding of GH3 enzymes, and to protein engineering of
228 isoprimeverose-producing enzymes for the production of novel oligosaccharides.
229
230 **Accession code**
231 The atomic coordinates and structural factors have been deposited in the Protein Data Bank under
232 accession codes 5YOT (apo) and 5YQS (isoprimeverose complex).

Acknowledgments

We thank the Beamline staff at BL44XU of SPring-8 for technical assistance during data collection. Synchrotron experiments were performed at SPring-8 under the approval of the Japan Synchrotron Radiation Research Institute under proposal numbers 2016A6657, 2016B6657, 2017A6767, and 2017B6767. This study was supported by the Japan Society for the Promotion of Science KAKENHI (Grant-in-Aid for Scientific Research B, Grant No. 18H02132).

Author contributions

T.M. and K.Y. designed the study. T.M., K.Y., and M.W. wrote the paper. T.M. cloned, expressed, purified, and measured the enzyme activity. M.W. crystallized the protein, collected, and processed the X-ray data, and refined and analyzed the structure. Y.N. assisted with the collection and processing of X-ray data. Z.F. assisted with refining and analyzing the structure. All authors commented on the manuscript.

Declaration of interests

The authors declare no competing interests.

References

Bauer, S., Vasu, P., Mort, A. J., and Somerville, C. R. (2005) Cloning, expression, and characterization of an oligoxyloglucan reducing end-specific xyloglucanobiohydrolase from *Aspergillus nidulans*. Carbohydr. Res. *340*, 2590–2597.

255 Carpita, N. C., and Gibeaut, D. M. (1993) Structure models of primary cell walls in flowering plants:
256 consistency of molecular structure with the physical properties of the walls during growth. Plant J. *3*,
257 1–30.
258 Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein
259 crystallography. Acta Cryst. *D50*, 760–763.
260 Eberhardt R. Y., Eddy S. R., Mistry J., Mitchell A. L., Potter S.C., Punta M., Qureshi M.,
261 Sangrador-Vegas A., Salazar G. A., Tate J., Bateman A. (2016) The Pfam protein families database:
262 towards a more sustainable future. Nucleic Acids Res. *44*, D279–285.
263 Edward, M., Bowman, Y. J., Dea, I. C., and Reid, J. S. (1988) A β-D-galactosidase from nasturtium
264 (*Tropaeolum majus* L.) cotyledons. Purification, properties, and demonstration that xyloglucan is the
265 natural substrate. J. Biol. Chem. *263*, 4333–4337.
266 Edwards, M., Dea, I. C., Bulpin, P. V., and Reid, J. S. (1986) Purification and properties of a novel
267 xyloglucan-specific endo-(1→4)-β-D-glucanase from germinated nasturtium seeds (*Tropaeolum*
268 *majus* L.). J. Biol. Chem. *261*, 9489–9494.
269 Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr.,
270 Sect. D: Biol. Crystallogr. *60,* 2126–2132.
271 Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A.,
272 Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J., and Punta, M. (2014) Pfam:
273 the protein families database. Nucleic Acids Res. *42*, D222–230.
274 Fox, J. D., and Robyt, J. F. (1991) Miniaturization of three carbohydrate analyses using a microsample
275 plate reader. Anal. Biochem. *195*, 93–96.

Fry S. C., York W. S., Albersheim P., Darvill A., Hayashi T., Joseleau J.-P., Kato Y., Lorences E. P., Maclachlan G. A., McNeil M., et al. (1993) An unambiguous nomenclature for xyloglucan-derived oligosaccharides. Physiol. Plant *89*, 1–3.

Grishutin, S. G., Gusakov, A. V., Markov, A. V., Ustinov, B. B., Semenova, M. V., and Sinitsym, A. P. (2004) Specific xyloglucanases as a new class of polysaccharide-degrading enzymes. Biochim. Biophys. Acta. 1674, 268–281.

Kato, Y., Matsushita, J., Kubodera, T., and Matsuda, K. (1985) A novel enzyme producing isoprimeverose from oligoxyloglucans of *Aspergillus oryzae*. J. Biochem. 97, 801–810.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 774–797.

Larsbrink, J., Thompson, A. J., Lundqvist, M., Gardner, J. G., Davies, G. J., and Brumer, H. (2014) A complex gene locus enables xyloglucan utilization in the model saprophyte *Cellvibrio japonicas*. Mol. Microbiol. *94*, 418–433.

Liisa, H., and Laura, M. L. (2016) Dali server update. Nucleic acids research. *44(W1)*, W351–355.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. *42*, D490–495.

Matsuzawa, T., Jo, T., Uchiyama, T., Manninen, J. A., Arakawa, T., Miyazaki, K., Fushinobu, S., and Yaoi, K. (2016a) Crystal structure and identification of a key amino acid for glucose tolerance, substrate specificity, and transglycosylation activity of metagenomic β-glucosidase Td2F2. FEBS J. 283, 2340–2353.

Matsuzawa, T., Kimura, N., Suenega, H., and Yaoi, K. (2016b) Screening, identification, and characterization of α-xylosidase from a soil metagenome. J. Biosci. Bioeng. *122*, 393–399.

298 Matsuzawa, T., Mitsuishi, Y., Kameyama, A., and Yaoi, K. (2016c) Identification of the gene encoding
299 isoprimeverose-producing oligoxyloglucan hydrolase in *Aspergillus oryzae*. J. Biol. Chem. *291*,
300 5080–5087.
301 Matsuzawa, T., Saito, Y., and Yaoi, K. (2014) Key amino acid residues for the endo-processive activity
302 of GH74 xyloglucanase. FEBS Lett. *588*, 1731–1738.
303 McCoy, A. J., Grosse-Kunstieve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J.
304 (2007) Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.
305 McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. (2011) Presenting your structures: the
306 CCP4mg molecular-graphics software. Acta Cryst. *D67*, 386-394.
307 Moracci, M., Cobucci Ponzano, B., Trincone, A., Fusco, S., De Rose, M., van Der Oost, J., Sensen, C. W.,
308 Charlebois, R. L., and Rossi, M. (2000) Identification and molecular characterization of the first
309 α-xylosidase from an Archaeon. J. Biol. Chem. *275*, 22082–22089.
310 Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D.,
311 Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal
312 structures. Acta Cryst. *D67*, 355–367.
313 Nakajima, M., Yoshida, R., Miyanaga, A., Abe, K., Takahashi, Y., Sugimoto, N., Toyoizumi, H., Nakai,
314 H., Kitaoka, M., and Taguchi, H. (2016) Functional and Structural Analysis of a β-Glucosidase
315 Involved in β-1,2-Glucan Metabolism in *Listeria innocua*. PLoS One. *11*, e0148870.
316 Okuyama, M., Mori, H., Chiba, S., and Kimura, A. (2004) Overexpression and characterization of two
317 unknown proteins, YicI and YihQ, originated from *Escherichia coli*. Protein Expr. Purif. *37*,
318 170–179.

319 Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode.
320 Methods Enzymol. *276,* 307–326.
321 Takeda, T., Furuta, Y., Awano, T., Mizuno, K., Mitsuishi, Y., and Hayashi, T. (2002) Suppression and
322 acceleration of cell elongation by integration of xyloglucans in pea stem segments. Proc. Natl. Acad.
323 Sic. USA *25*, 9055–9060.
324 Tuomivaara, S. T., Yaoi, K., O'Neill, M. A., and York, W. S. (2015) Generation and structural validation
325 of a library of diverse xyloglucan-derived oligosaccharides, including an update on xyloglucan
326 nomenclature. Carbohydr. Res. *30*, 56–66.
327 Varghese, J. N, Hrmova, M., and Fincher, G. B. (1999) Three-dimensional structure of a barley
328 β-D-glucan exohydrolase, a family 3 glycosyl hydrolase. Structure *7*, 179–190.
329 Yaoi, K., Hiyoshi, A., and Mitsuishi, Y. (2007) Screening, purification and characterization of a
330 prokaryotic isoprimeverose-producing oligoxyloglucan hydrolase from *Oerskovia* sp. Y1. J. Appl.
331 Glycosci. 54, 91–94.
332 Yaoi, K., and Mitsuishi, Y. (2002) Purification, characterization, cloning, and expression of a novel
333 xyloglucan-specific glycosidase, oligoxyloglucan reducing end-specific cellobiohydrolase. J. Biol.
334 Chem. *277*, 48276–48281.
335 Yaoi, K., and Mitsuishi, Y. (2004) Purification, characterization, cDNA cloning, and expression of a
336 xyloglucan endo-glucanase from *Geotrichum* sp. M128. FEBS Lett. *560*, 45–50.
337 Yaoi, K., and Miyazaki, K. (2012) Cloning and expression of isoprimeverose-producing oligoxyloglucan
338 hydrolase from actinomycetes species, *Oerskovia* sp. Y1. J. Appl. Glycosci. *59*, 83–88.
339 Yaoi, K., Nakai, T., Kameda, Y., Hiyoshi, A., and Mitsuishi, Y. (2005) Cloning and characterization of
340 two xyloglucanases from *Paenibacillus* sp. strain KM21. Appl. Environ. Microbiol. *71*, 7670–7678.

York, W.S., Harvey, L.K., Guillen, R., Albersheim, P., and Darvill, A.G. (1993) Structural analysis of tamarind seed xyloglucan oligosaccharides using β-galactosidase digestion and spectroscopic methods. Carbohydr. Res. *248*, 285–301.

Zabotina, O. A. (2012) Xyloglucan and its biosynthesis. Front. Plant Sci. *3*, 134.

Figure legends

Figure 1. Overall structure of *Aspergillus oryzae* isoprimeverose-producing enzyme (IpeA). (A) Homodimeric structures (chains A and B) of IpeA in two views, shown as ribbon diagrams. Arrow and ellipse show 2-fold axis of dimer. (B) A monomeric structure (residues 26–779) of IpeA is composed of three distinct domains shown in a ribbon diagram. The domains are indicated by different colors: green (TIM-barrel, 48–371), purple (α/β/α sandwich, 411–660), and yellow (fibronectin-like domain, 698–767).

Figure 2. Structural comparisons of IpeA and other glycoside hydrolase family 3 (GH3) β-glucosidases. A Dali search for structural similarities to the monomer structure of IpeA identified three GH3 β-glucosidases: β-1,2-glucooligosaccharide-specific GH3 β-glucosidase Lin1840r from *Listeria innocua* (Z = 45.1, rmsd = 2.0 Å, PDB 4ZOE), a GH3 β-glucosidase JMB19063 isolated from compost (Z = 44.0, rmsd = 1.8 Å, PDB 3U48), and a barley β-D-glucan exohydrolase isoenzyme ExoI from young seedlings (Z = 43.3, rmsd = 1.9 Å, PDB 1EX1). The structures are colored as follows: IpeA (pink), Lin1840r (coral) JMB19603 (ice blue), and ExoI (cyan).

Figure 3. Active site comparisons of IpeA and β-glucosidase JMB19603. Superposition of the IpeA monomer (pink) with isoprimeverose (green cylinder model) (A) and JMB19603 (ice blue) with D-glucose (green cylinder model) (B) are shown as molecular surface models. (C, D) Insets show magnifications of both active sites. The amino acids of IpeA (pink stick with black and red labels) and JMB19603 (ice blue stick with black labels) are in complex with isoprimeverose and D-glucose (green sticks), respectively. The hydrogen bonds between the IpeA side chains (pink stick) and isoprimeverose (green stick with hydroxyl or carboxyl group numbers), and JMB19603 (ice blue stick) and D-glucose (green stick with hydroxyl or carboxyl group numbers) are represented as dashed lines. Oxygen, nitrogen, and sulfur atoms are shown in red, blue, and yellow, respectively. Both *Fo-Fc* electron density maps are contoured at 2 σ in blue around isoprimeverose and D-glucose. These figures are drawn in the same orientation as Figures 3A and B. (E, F) Schematic diagrams showing the interaction of IpeA with isoprimeverose (E) and JMB19603 with D-glucose (F). Dashed lines represent hydrogen bonds, with the interatomic distance in angstroms.

Figure 4. Superimposed models of the active site of IpeA and other GH3 β-glucosidases. The stick models of the active sites are colored as follows: IpeA (pink) with isoprimeverose (green stick), Lin1840r (tan), JMB19603 (ice blue), and ExoI (cyan). The labels are indicated in plain black (IpeA) and italic text (other GH3 β-glucosidases).

Figure 5. Putative positive subsites of IpeA with a xyloglucan oligosaccharide model (XXXG, Xyl_3Glc_4). (A) IpeA is shown as the surface model (white). (B) The putative residues around the positive subsite are depicted in red (Glu67), magenta (Tyr92 and Tyr268), blue (Arg306), and purple (Trp515).

384 The model of XXXG is shown as a cylinder model with carbon (green) and oxygen (red) atoms. The

385 figure is magnified in the same orientation as Figure 5A.

386

Table 1. Data collection and refinement statistics of the IpeA crystal structure

	Apo	Isoprimeverose-complex
Data Collection		
Space group	$P2_1$	$P2_1$
Cell dimensions a, b, c (Å)	70.14, 130.45, 94.83	69.78, 129.91, 95.16
α, β, γ (°)	90, 94.46, 90	90, 94.44, 90
Wavelength	0.9000	0.9000
Resolution (Å)	50.00–1.98 (2.03–1.98)*	50.00–2.40 (2.44–2.40)*
R_{merge}[a] (%)	9.7 (65.1)	8.5 (23.7)
$I/\sigma I$	25.5 (3.8)	20.7 (5.4)
Completeness (%)	93.9 (93.2)	90.6 (81.4)
Redundancy	3.9 (3.7)	3.0 (2.2)
Refinement		
Resolution (Å)	30.0–1.98	30.0–2.40
No. reflections	109,249	59,761
R_{work}[b] / R_{free}[c]	0.168/0.214	0.152/0.207
No. atoms		
Protein	11,702	11,670
Ligands	386	434
Solvent	814	552
B factors		
Protein (Å²)	27.6	26.0
Ligands (Å²)	51.9	51.5
Solvent (Å²)	29.8	21.6
r.m.s. deviations from ideals		
Bond lengths (Å)	0.01	0.01
Bond angles (°)	1.7	1.6

*Outer shells (apo; 2.03-1.98 Å, isoprimeverose-complex; 2.44-2.40 Å).

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i \mid I_i(hkl) - \langle I(hkl)\rangle \mid \Sigma_{hkl}\Sigma I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th measurement of reflection hkl, including symmetry-related reflections, and $\langle I(hkl)\rangle$ is their average. [b] $R_{work} = \Sigma_h\Sigma_i \mid\mid F_o\mid - \mid F_c\mid\mid / \Sigma \mid F_o\mid$. [c] R_{free} is R_{work} for approximately 5 % of the reflections that were excluded from the refinement.

395

396 **Table 2. Substrate specificities of *ipeA* mutants**

	Activity, μmol/min/mg (Relative activity, %)			
	5 mM XG	5 mM XX	5 mM XXXG	5 mM LG
Wild-type	253 ± 2 (100%)	332 ± 5 (100%)	519 ± 2 (100%)	n.d.
Gln58Ala	n.d.	2.85 ± 0.05 (0.86%)	12.4 ± 0.6 (2.38%)	n.d.
Tyr89Ala	n.d.	7.74 ± 0.18 (2.33%)	56.5 ± 0.3 (10.9%)	n.d.
Tyr89Trp	77.4 ± 2.5 (30.6%)	139 ± 1 (41.7%)	229 ± 4 (44.1%)	n.d.
Tyr89Phe	216 ± 2 (85.2%)	292 ± 3 (87.8%)	416 ± 3 (80.3%)	n.d.
Tyr268Ala	n.d.	1.27 ± 0.30 (0.38%)	n.d.	n.d.

397 n.d.: not detected

398

399

400

401

402 **Table 3. Kinetic analysis of *ipeA* subsite -1' mutants for a reduced XXXG substrate**

	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}mM^{-1}$)
Wild-type	0.649 ± 0.051	678 ± 18	1045
Gln58Ala	1.82 ± 0.24	17.9 ± 0.8	9.85
Tyr89Ala	2.53 ± 0.13	69.4 ± 1.4	27.4
Tyr89Trp	0.460 ± 0.061	301 ± 13	655
Tyr89Phe	0.611 ± 0.041	531 ± 12	869

403

404

405

406 **STAR Methods**

407 **Method Details**

408 **Materials**

409 Xyloglucan oligosaccharides were prepared as previously described (Matsuzawa et al., 2016c; Takeda et

410 al., 2002; Yaoi and Mitsuishi, 2002; Yaoi et al., 2007).

411

412 **Expression and purification of IpeA**

413 IpeA was expressed in *Pichia pastoris* strain X-33 and purified using a Ni^{2+} affinity column (HisTrap FF;

414 GE Healthcare, Buckinghamshire, UK) as described previously (Matsuzawa et al., 2016c). Purified IpeA

415 was treated with endoglycosidase H (endo H; New England Biolabs, Ipswich, MA, USA) as follows. An

416 8-mL reaction mixture containing purified IpeA, 10 mM sodium acetate buffer (pH 6.0), and 10,000 units

417 of endo H was incubated at 30°C for 1 h. To remove endo H and release sugar chains, endo H-treated

418 IpeA was purified using a Ni^{2+} affinity column as previously described.

419

420 **Protein crystallization**

421 The purified endo H-treated IpeA was concentrated to 7.3 mg/mL using a Vivaspin 20 concentrator

422 (10,000 MWCO, Sartorius AG, Goettingen, Germany). Initial crystallization screening was performed

423 manually using an Index HT kit (Hampton Research, Aliso Viejo, CA, USA) by the sitting-drop vapor

424 diffusion method at 293 K. Each drop contained 0.3 µL protein and 0.3 µL reservoir solution, and was

425 equilibrated with 60 µL reservoir solution. After 2 weeks, small crystals were obtained under several

426 conditions, of which No. 54 (0.05 M calcium chloride, 0.1 M Bis-Tris pH 6.5, and 30% [v/v] PEG MME

427 550) was the most promising, as suggested by Index HT. The best quality crystals of the enzyme were

obtained under the condition using 0.05 M calcium chloride, 0.15 M Bis-Tris pH 6.5, and 26% (v/v) PEG MME 550. The crystals of isoprimeverose complex were obtained by soaking with 50 mM isoprimeverose.

X-ray crystallography

Crystals of the apo form and isoprimeverose complex of IpeA were cryoprotected with 20% (v/v) glycerol, and subsequently flash-cooled in liquid nitrogen. All data sets for the both crystals were collected on a BL44XU instrument (Hyogo, Japan) with a Rayonix MX300HE detector under a cryostream at 90 K, and the data were processed and scaled with the HKL-2000 program suite (Otwinowski and Minor, 1997). Data sets for the apo form and isoprimeverose complex of IpeA were collected at a wavelength of 0.9 Å. General data handling was carried out with the CCP4 package (Collaborative Computational Project, Number 4, 1994), and the structures were solved by molecular replacement using Phaser (McCoy et al., 2007). Manual adjustment of the models was carried out with COOT (Emsley and Cowtan, 2004), and refinement using REFMAC5 (Murshudov et al., 2011). Water molecules were added to the models by manual inspection of the 2*Fo-Fc* and *Fo-Fc* maps. Ramachandran plots for the apo form and isoprimeverose complex of IpeA models showed 97.2 and 97.0% of the residues, respectively, to be in the most favored regions, with the remaining 2.6 and 2.9% of the residues in additional allowed regions. Data collection and refinement statistics are shown in Table 1. Figures were created using CCP4MG (McNicholas et al., 2011).

Mutation at subsite −1'

IpeA mutants (Gln58Ala, Tyr89Ala, Tyr89Trp, Tyr89Phe, and Tyr268Ala) were constructed as described below. Polymerase chain reaction (PCR) was performed using pGAPZα A-IpeA plasmid (Matsuzawa et al., 2016c) as a template, primers (primers 5'-AATCAATTGAGACATCTTCTCCTCGAGC-3' and 5'-ATGTCTCAATTGATTGCTGGCGACATCACCAATTGGATGAATG-3' for Gln58Ala; primers 5'-GAACATTCCACCCCTCATTTTCGTAC-3' and 5'-AGGGGTGGAATGTTCGCTGTCGGATATCCGGTGCCTTGG-3', 5'-GAACATTCCACCCCTCATTTTCGTAC-3' and 5'-AGGGGTGGAATGTTCTGGGTCGGATATCCGGTGCCTTGG-3', and 5'-GAACATTCCACCCCTCATTTTCGTAC-3' and 5'-AGGGGTGGAATGTTCTTTGTCGGATATCCGGTGCCTTGG-3' for Tyr89Ala, Tyr89Trp, and Tyr89Phe, respectively; and primers 5'-CGCACTCATGATACTCCACGC-3' and 5'-AGTATCATGAGTGCGGCTCACTCATACGACGGTATCCCCG-3' for Tyr268Ala), and PrimeSTAR Max DNA polymerase (TaKaRa Bio, Inc., Shiga, Japan). PCR products were treated with *Dpn*I to degrade the template. Amplified DNA fragments were ligated using In-Fusion HD (TaKaRa Bio, Inc.). pGAPZα A-*ipeA* mutated plasmid (Gln58Ala, Tyr89Ala, Tyr89Trp, Tyr89Phe, or Tyr268Ala) was digested with *Avr*II and integrated into *P. pastoris* strain X-33. Mutated enzymes were expressed and purified as described previously (Matsuzawa et al., 2016c).

Enzymatic activity of IpeA mutants

The substrate specificities of IpeA mutants were determined as described below. Reaction mixture (50 μL) containing 5 mM substrate (XG, XX, LG, or XXXG), 50 mM sodium phosphate buffer (pH 4.5), and 0.05 μg of purified enzyme was incubated at 60℃ for 5 min. To stop the reaction, the reaction was

471 incubated at 98°C for 10 min. The released isoprimeverose or D-glucose was measured using

472 high-performance liquid chromatography (HPLC) as described previously (Matsuzawa et al., 2016c).

473

474 **Kinetic analysis of IpeA mutants**

475 Kinetic parameters (K_m, k_{cat}, and k_{cat}/K_m) of IpeA mutants mutated at subsite −1' were measured as

476 follows. The total reaction volume was 20 μL. The reaction mixture containing reduced XXXG substrate

477 (XXXGol: 0.125–4 mM for wild-type, Tyr89Trp, and Tyr89Phe mutants; 0.3125–10 mM for Gln58Ala

478 andTyr89Ala mutants), purified enzyme (0.5 ng for wild-type, Tyr89Trp, and Tyr89Phe mutants; 5 ng for

479 Gln58Ala and Tyr89Ala mutants), and 50 mM sodium acetate buffer (pH 4.5) was incubated at 60°C for

480 5 min. To inactivate the enzymes, the reaction mixture was incubated at 98°C for 10 min. The reducing

481 sugar produced was measured using the bicinchoninate assay (Fox and Robyt, 1991). A standard curve

482 was constructed using isoprimeverose. The kinetic constants (K_m and V_{max}) were calculated by non-linear

483 regression of the Michaelis–Menten equation using GraphPad Prism version 5.0 software (GraphPad

484 Software, Inc., La Jolla, CA, USA).

485

Figure 1. Matsuzawa et al.

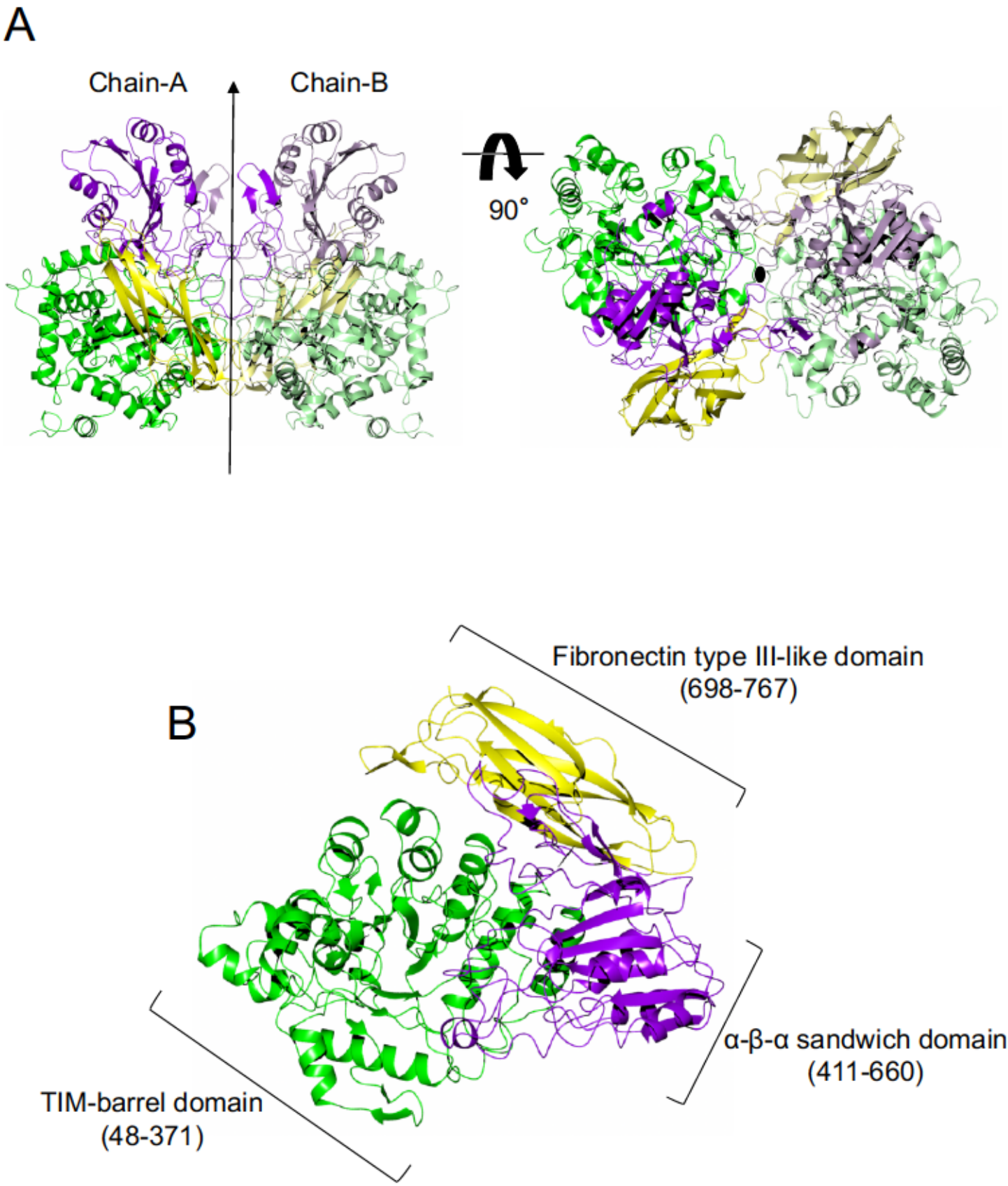


Figure 2. Matsuzawa et al.

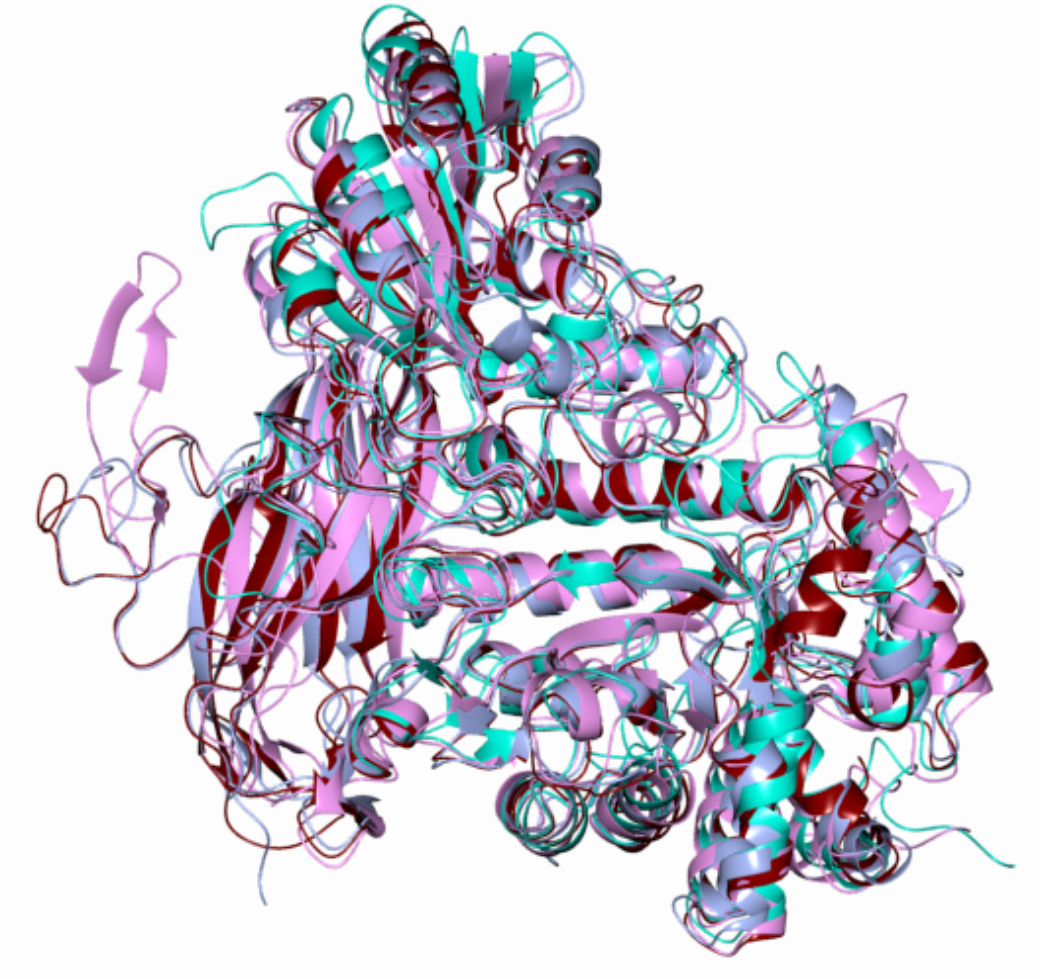


Figure 3. Matsuzawa et al.

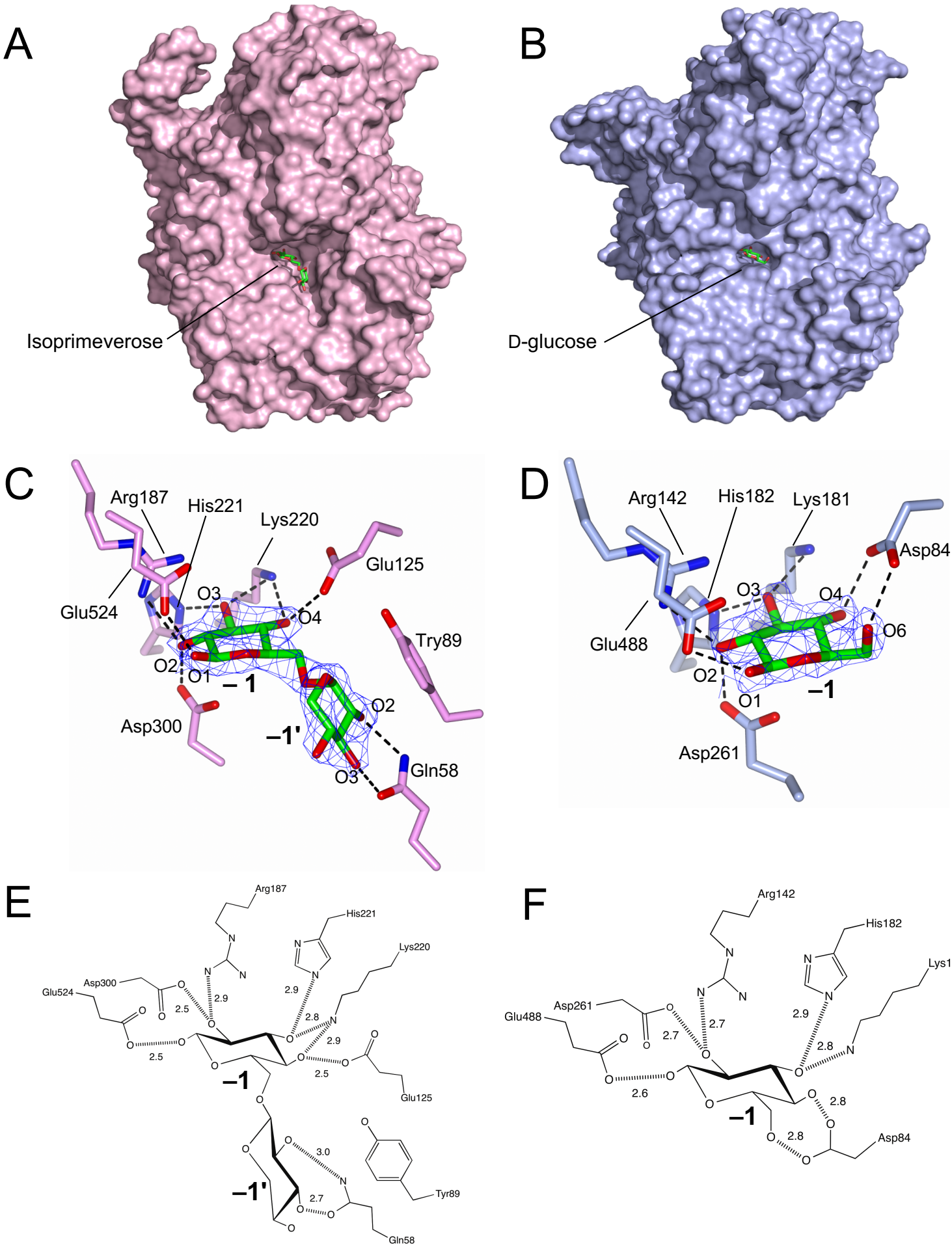


Figure 4. Matsuzawa et al.

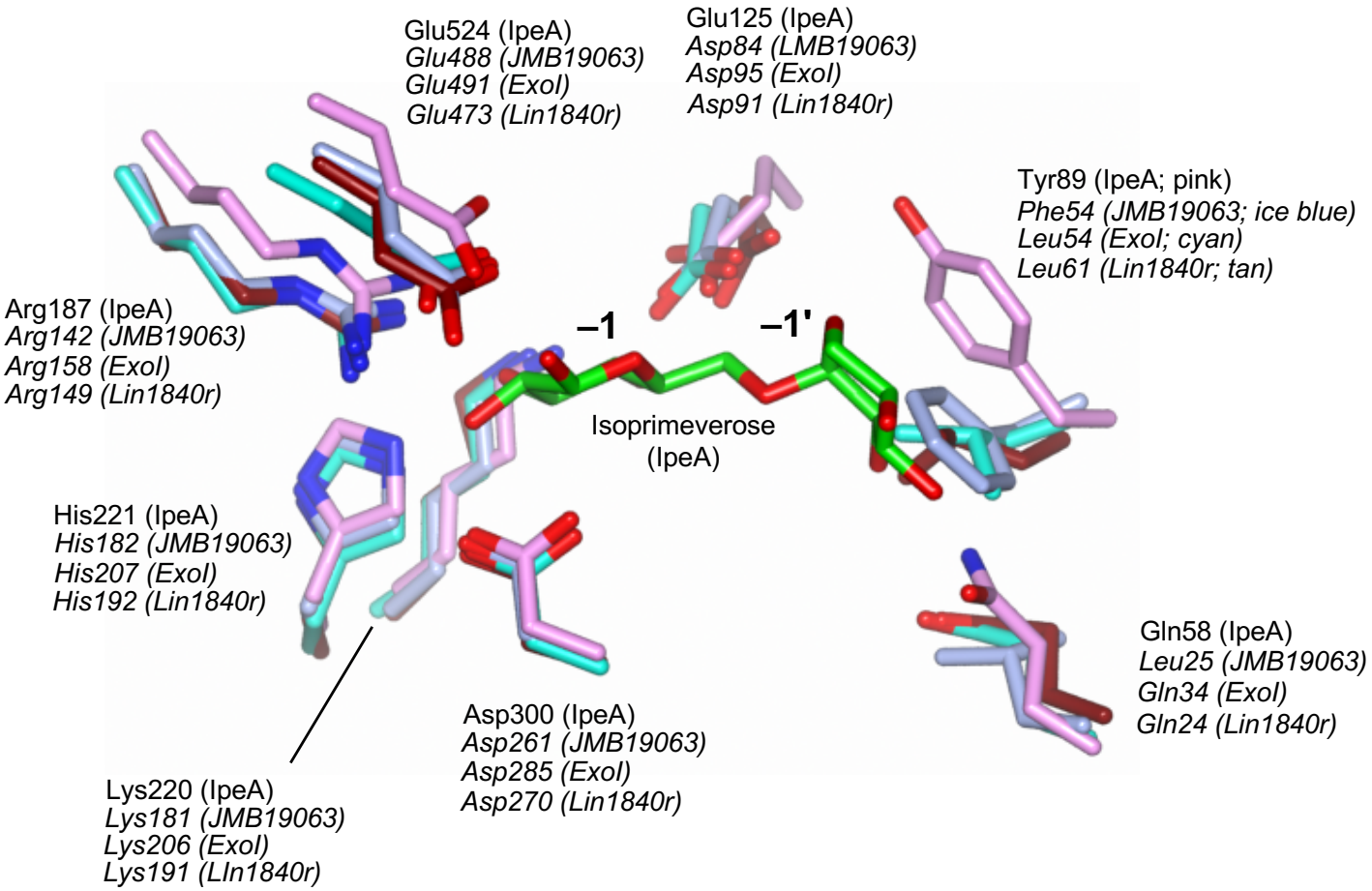
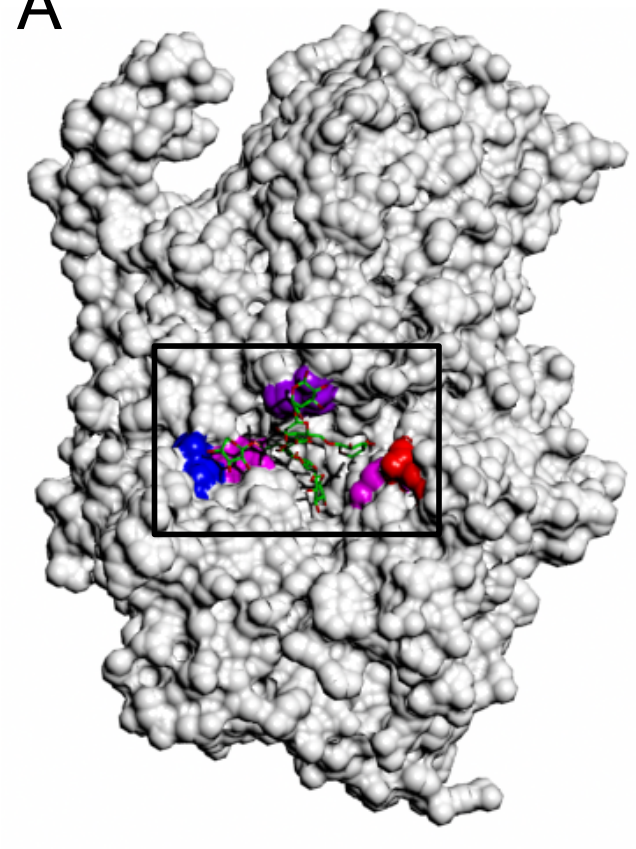


Figure 5. Matsuzawa et al.

A



B